



# A comparison of regionalisation methods for catchment model parameters

J. Parajka, R. Merz, G. Blöschl

## ► To cite this version:

J. Parajka, R. Merz, G. Blöschl. A comparison of regionalisation methods for catchment model parameters. Hydrology and Earth System Sciences Discussions, 2005, 9 (3), pp.157-171. hal-00304814

**HAL Id: hal-00304814**

**<https://hal.science/hal-00304814>**

Submitted on 9 Aug 2005

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A comparison of regionalisation methods for catchment model parameters

J. Parajka<sup>1,\*</sup>, R. Merz<sup>1</sup>, and G. Blöschl<sup>1</sup>

<sup>1</sup>Institute for Hydraulic and Water Resources Engineering, Vienna University of Technology, Austria

\*on leave from: Institute of Hydrology, Slovak Academy of Sciences, Bratislava, Slovakia

Received: 3 March 2005 – Published in Hydrology and Earth System Sciences Discussions: 12 April 2005

Revised: 17 June 2005 – Accepted: 1 August 2005 – Published: 9 August 2005

**Abstract.** In this study we examine the relative performance of a range of methods for transposing catchment model parameters to ungauged catchments. We calibrate 11 parameters of a semi-distributed conceptual rainfall-runoff model to daily runoff and snow cover data of 320 Austrian catchments in the period 1987–1997 and verify the model for the period 1976–1986. We evaluate the predictive accuracy of the regionalisation methods by jack-knife cross-validation against daily runoff and snow cover data. The results indicate that two methods perform best. The first is a kriging approach where the model parameters are regionalised independently from each other based on their spatial correlation. The second is a similarity approach where the complete set of model parameters is transposed from a donor catchment that is most similar in terms of its physiographic attributes (mean catchment elevation, stream network density, lake index, areal proportion of porous aquifers, land use, soils and geology). For the calibration period, the median Nash-Sutcliffe model efficiency *ME* of daily runoff is 0.67 for both methods as compared to *ME*=0.72 for the at-site simulations. For the verification period, the corresponding efficiencies are 0.62 and 0.66. All regionalisation methods perform similar in terms of simulating snow cover.

## 1 Introduction

Predicting hydrological variables in ungauged catchments has been singled out as one of the major issues in the hydrological sciences (Sivapalan et al., 2003). Predictions are particularly difficult to make in alpine regions where data are sparse and the spatial variability of the hydrological environment is enormous. Transferring information from neighbouring catchments to the catchment of interest is gener-

ally accomplished by hydrological regionalisation methods (Blöschl and Sivapalan, 1995). Numerous regionalisation methods have been proposed in the literature for the case of catchment model parameters (Blöschl, 2005). Among the most widely used techniques are regressions between the model parameters and physiographic catchment attributes. Typically, linear multiple regressions are used where each model parameter is estimated independently from the others (e.g. Post and Jakeman, 1996, 1999; Sefton and Howarth, 1998). The regressions are not always straightforward to interpret. In a comparative study of 331 catchments in Australia, Peel et al. (2000), for example, did not find the parameters of the SYMHID model significantly correlated to the catchment attributes. Seibert (1999) related the model parameters of the HBV model to the attributes of 11 Swedish catchments within the NOPEX area. Relationships between lake percentage and soil parameters found by Seibert (1999) could not be explained by hydrologic reasoning while relationships between forest percentage and snow parameters supported the process basis of the model. Similar conclusions were drawn by Kokkonen et al. (2003). They used the IHACRES model with 6 parameters and found that high significance of regressions does not guarantee a set of parameters with a good predictive power. Care must hence be taken when interpreting the physical meaning of parameter-descriptor relationships found by regressions.

The regression method is the most widely used regionalisation technique but alternative methods are in use. Vandewiele and Elias (1995) examined two methods based on spatial proximity, the kriging method and the use of model parameter values from a few neighbouring catchments in a Belgian case study. They found that the kriging approach provided a significantly better model performance than the nearest neighbour approach although the model performance for some of the catchments was rather poor. The question of whether or not homogeneous catchments tend to occur in close proximity to each other has been the subject of

---

Correspondence to: J. Parajka  
(parajka@hydro.tuwien.ac.at)

significant debate over the years. Shu and Burn (2003) suggested that geographically close catchments are not necessarily homogeneous in terms of hydrological response. In their case study in Great Britain, homogeneous spatial clustering patterns of the regional flood frequency distribution were found within a 62.5 km radius from a local clustering centre. Burn and Boorman (1993) assigned donor catchments based on a similarity measure of physiographic catchment characteristics. In this method, the catchment characteristics are similar to those of the regression approach but the regionalisation model structure is different as no assumption of linearity is made. Also, the complete set of model parameters is usually transposed from one or more donor catchments to the catchment of interest in this approach, while in the regression case, the parameters are usually regionalised independently from each other. Along similar lines, Campbell and Bates (2001) used a regional link function to estimate the parameters of a quasi-distributed, non-linear flood event model for 39 watersheds in Australia with good accuracy. Fernandez et al. (2000) proposed a regional calibration approach that involves a concurrent calibration of the model parameters and the relationships between model parameters and catchment attributes at many sites in a region. This approach has led to nearly perfect regional relationships between model parameters and catchment characteristics, however, these relationships did not improve the streamflow predictions at ungauged sites. A similar approach was applied in Hlavčová et al. (2000) and Szolgay et al. (2003), where they intended to find regionally valid parameters of a monthly water balance model. They jointly calibrated a model using multiobjective calibration, where the catchments were pooled together using cluster analysis of selected physiographic catchment attributes.

Merz and Blöschl (2004) examined the performance of various methods of regionalising the parameters of a conceptual catchment model in 308 Austrian catchments. They concluded that the methods based on spatial proximity performed better than those based on physiographic catchment attributes. The present paper builds on their analysis and examines the relative performance of regionalisation methods. The present paper goes beyond Merz and Blöschl (2004) in three important aspects.

First, Merz and Blöschl (2004) used a lumped catchment model. In an alpine country such as Austria there may be merits in allowing different model states in different elevations of the catchment to improve the overall predictive performance. In this paper we hence subdivide each catchment into elevation zones of 200 m. Second, even though Merz and Blöschl (2004) tested the robustness of model parameters in a comprehensive way, further gains in robustness may be obtained by a multi-objective calibration where response data in addition to runoff are used. In this paper we hence use both runoff data and snow cover data to calibrate and validate the model. Third, Merz and Blöschl (2004) found that the regressions between model parameters and catchment at-

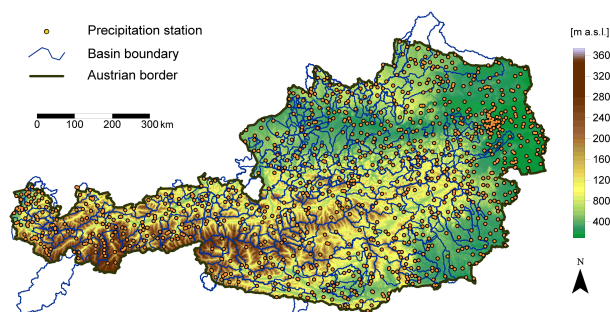
tributes performed not as well as other methods but it was not clear whether this was due to the catchment attributes being poor hydrological indicators at the regional scale or due to problems with the linearity assumption of the multiple linear regressions used. In this paper we hence examine alternative methods that use catchment attributes and are based on similarity measures.

This paper is organised as follows. We first provide a brief description of the dataset and give an overview of the hydrologic model and the calibration procedure. In the following section we describe the regionalisation approaches and the methodology used for cross-validation. We then present the results in terms of model performance of the different regionalisation methods and discuss the main findings of the paper.

## 2 Data

This study was carried out in Austria using data from the period 1976–1997. Austria is flat or undulating in the east and north, and alpine in the west and south. Elevations range from 115 m a.s.l. to 3797 m a.s.l.. Mean annual precipitation is less than 400 mm/year in the east and almost 3000 mm/year in the west. Land use is mainly agricultural in the lowlands and forest in the medium elevation ranges. Alpine vegetation and rocks prevail in the highest catchments. The dataset used in this study includes measurements of daily precipitation and snow depths at 1091 stations and daily air temperature at 212 climatic stations. To calibrate and verify a catchment model, daily runoff data from 320 gauged catchments were used with areas ranging from 10 km<sup>2</sup> to 9770 km<sup>2</sup> and a median of 196 km<sup>2</sup>. 97 of these catchments range in area between 10 and 100 km<sup>2</sup>, 106 catchments between 100 and 300 km<sup>2</sup>, 64 catchments between 300 and 1000 km<sup>2</sup> and 55 catchments have areas of more than 1000 km<sup>2</sup>. In preliminary analyses we carefully screened the runoff data for errors and removed all stations with significant anthropogenic effects. We also removed stations where we were not able to close the long term water balance. The spatial distribution of the climate stations and the boundaries of the gauged catchments are shown in Fig. 1.

The inputs to the water balance model were prepared in two steps. First, the daily values of precipitation, snow depth and air temperature were spatially interpolated by methods that use elevation as auxiliary information. External drift kriging was used for precipitation and snow depths, and the least-squares trend prediction method was used for air temperatures (Pebesma, 2001). The spatial distribution of potential evapotranspiration was estimated by a modified Blaney-Criddle method (Parajka et al., 2003) using daily air temperature and potential sunshine duration calculated by the Solei-32 model (Mészáros et al., 2002) that incorporates shading by surrounding terrain. In a second step, a digital elevation model with a 1 × 1 km grid resolution was used for deriving 200 m elevation zones in each catchment. Time-series



**Fig. 1.** Topography (m a.s.l.) of Austria and boundaries of the gauged catchments used in this paper. The dots show the locations of stations with precipitation and snow depth measurements.

of daily precipitation, air temperature, potential evaporation and snow depth were then extracted for each of the elevation zones to be used in the water balance simulations.

For testing different regionalisation approaches, we derived a range of physiographic catchment attributes. The topographic attributes were catchment area, catchment average and coefficient of variation of topographic elevation, average and the coefficient of variation of topographic slope and the minimum and maximum of the topographic wetness index of Beven and Kirkby (1979). Stream network density was calculated from the digital stream network map at the 1:50 000 scale for each catchment. The FARL (flood attenuation by reservoirs and lakes) lake index was calculated according to the Flood Estimation Handbook (1999). The attributes related to precipitation were the catchment average of long term mean annual precipitation, the average of the long term mean of maximum annual daily summer and winter precipitation for which the record length ranged from 45 to 97 years as well as the long term mean of maximum annual 1 hourly rainfall intensity from shorter records. The boundaries of porous aquifers were taken from the Hydrographic Yearbook (HZB, 2000) from which the areal proportion of porous aquifers in each catchment was estimated. Digital maps of land use (Ecker et al., 1995), regional soil types (based on the FAO map, see ÖBG, 2001) and the main geological formations (Geologische Bundesanstalt, 1998) were also used to derive the respective areal proportions in each catchment.

### 3 Model structure and model calibration

The model used in this paper is a semi-distributed conceptual rainfall-runoff model, following the structure of the HBV model (Bergström, 1976 and Lindström et al., 1997). The model runs on a daily time step and consists of a snow routine, a soil moisture routine and a flow routing routine. The snow routine represents snow accumulation and melt by a simple degree-day concept, using degree-day factor  $DDF$

**Table 1.** Model parameters and a priori distributions.  $u$  and  $v$  are the parameters of the Beta distribution (Eq. 8),  $p_l$  and  $p_u$  are the lower and upper bounds of the parameter space and  $p_{\max}$  is the parameter value at which the Beta distribution is at a maximum (Eq. 6).

Parameter name	Model part	$p_l$	$p_u$	$u$	$v$	$p_{\max}$
SCF [–]	Snow	1.0	1.5	1.2	4.0	1.03
DDF [mm/°C day]	Snow	0.0	5.0	2.0	4.0	1.25
LP/FC [–]	Soil	0.0	1.0	4.0	1.2	0.94
FC [mm]	Soil	0.0	600.0	1.1	1.5	100.2
$\beta$ [–]	Soil	0.0	20.0	1.1	1.5	3.4
$K_0$ [days]	Runoff	0.0	2.0	2.0	4.0	0.5
$K_1$ [days]	Runoff	2.0	30.0	2.0	4.0	9.0
$K_2$ [days]	Runoff	30.0	180.0	1.05	1.05	105.0
$C_P$ [mm/day]	Runoff	0.0	8.0	2.0	4.0	2.0
$C_R$ [days <sup>2</sup> /mm]	Runoff	0.0	50.0	1.05	1.05	25.0
$LS_{UZ}$ [mm]	Runoff	1.0	100.0	3.0	3.0	50.5

and melt temperature  $T_M$ . Catch deficit of the precipitation gauges during snowfall is corrected by a snow correction factor,  $SCF$ . A threshold temperature interval  $T_R - T_S$  is used to distinguish between rainfall, snowfall and a mix of rain and snow. The soil moisture routine represents runoff generation and changes in the soil moisture state of the catchment and involves three parameters: the maximum soil moisture storage  $FC$ , a parameter representing the soil moisture state above which evaporation is at its potential rate, termed the limit for potential evaporation  $LP$ , and a parameter in the non-linear function relating runoff generation to the soil moisture state, termed the non-linearity parameter  $\beta$ . Runoff routing on the hillslopes is represented by an upper and a lower soil reservoir. Excess rainfall enters the upper zone reservoir and leaves this reservoir through three paths, outflow from the reservoir based on a fast storage coefficient  $K_1$ ; percolation to the lower zone with a constant percolation rate  $C_P$ ; and, if a threshold of the storage state  $LS_{UZ}$  is exceeded, through an additional outlet based on a very fast storage coefficient  $K_0$ . Water leaves the lower zone based on a slow storage coefficient  $K_2$ . The outflow from both reservoirs is then routed by a triangular transfer function representing runoff routing in the streams, where the base of transfer function is estimated with the scaling of the outflow by the  $C_R$  parameter. The model concept is similar to that presented in Merz and Blöschl (2004). The difference is that in this study we used a semi-distributed model structure of 200 m elevation zones while the model of Merz and Blöschl (2004) was spatially lumped.

The model was run for all 320 gauged catchments in Austria. Daily inputs (precipitation, air temperature and potential evapotranspiration) were allowed to vary with elevation within a catchment, and the soil moisture accounting and



snow accounting was performed independently in each elevation zone. However, the same model parameters were assumed to apply to all elevation zones of a catchment. From a total of 14 model parameters, 3 parameters were preset ( $T_R=2^\circ\text{C}$ ,  $T_S=0^\circ\text{C}$ ,  $T_M=0^\circ\text{C}$ ) and 11 parameters (Table 1) were estimated by automatic model calibration. We used the shuffled complex evolution (SCE-UA) scheme of Duan et al. (1992) to calibrate the model parameters to observed runoff and snow cover. The objective function ( $Z_C$ ) used in the calibration involves three parts which are related to runoff ( $Z_Q$ ), snow cover ( $Z_S$ ) and a priori information about the distribution of each model parameter ( $Z_P$ ).  $Z_C$  is the weighted mean of these parts:

$$Z_C = w_1 \cdot Z_Q + w_2 \cdot Z_S + w_3 \cdot Z_P, \quad (1)$$

where the weights were obtained in test simulations as  $w_1=0.6$ ,  $w_2=0.1$  and  $w_3=0.3$ . These consisted of sensitivity analyses that showed that the model results were only moderately sensitive to the choice of weights.

The runoff objective function  $Z_Q$  combines the Nash-Sutcliffe coefficient ( $ME$ ) and the relative volume error ( $VE$ ):

$$Z_Q = (1 - ME) + w_4 \cdot VE, \quad (2)$$

where

$$ME = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2}, \quad (3)$$

$$VE = \frac{\sum_{i=1}^n Q_{sim,i} - \sum_{i=1}^n Q_{obs,i}}{\sum_{i=1}^n Q_{obs,i}}. \quad (4)$$

$Q_{sim,i}$  is the simulated streamflow on day  $i$ ,  $Q_{obs,i}$  is the observed streamflow,  $\overline{Q_{obs}}$  is the average of the observed streamflow over the calibration (or verification) period of  $n$  days, and the weight  $w_4$  was found from test simulations as  $w_4=0.1$ .

The snow objective function  $Z_S$  used observed and simulated snow coverage. Observed snow coverage was estimated from daily grid maps constructed from the observed snow depth data. If the catchment zone average of snow depth was greater than 0.5 mm than the zone was considered as snow covered, otherwise as snow free. Simulated snow coverage was derived from the snow water equivalent simulated by the model where a zone was considered snow covered if the water equivalent was greater than 0.1 mm, otherwise it was considered snow free. Snow simulations on a particular day were considered to be poor if the absolute difference between simulated and observed snow coverage was greater than 50% of the catchment area. The snow objective function  $Z_S$  was then defined as the ratio of the number of days with poor

snow cover simulation ( $n_{ps}$ ) to the total number of days in the simulation period:

$$Z_S = \frac{n_{ps}}{n} \quad (5)$$

The third term,  $Z_P$ , allows inclusion of an expert estimate about the a priori distribution of each parameter. In calibration procedures, the parameter values are usually bounded between two limits (Duan et al., 1992) and otherwise no a priori assumptions are made about the parameters. This implies that the a priori distribution of the parameters is a uniform distribution. We believe that it is possible to make a more informed guess about the shape of the a priori distribution of the model parameters and introduced a penalty function,  $Z_P$ , based on an a priori Beta distribution for each parameter:

$$Z_P = \sum_{j=1}^k \frac{f_{\max,j} - f_j \left( \frac{p_j - p_{l,j}}{p_{u,j} - p_{l,j}} \right)}{f_{\max,j}} \quad (6)$$

$$f_{\max,j} = f_j \left( \frac{p_{\max,j} - p_{l,j}}{p_{u,j} - p_{l,j}} \right) \quad (7)$$

where  $p_j$  is the model parameter  $j$  to be calibrated,  $p_l$  and  $p_u$  are the lower and upper bounds of the parameter space, respectively,  $p_{\max}$  is the parameter value corresponding to the mode of the a priori distribution and  $k$  is the number of parameters to be calibrated. The probability density function of the Beta distribution  $f$  is:

$$f(x|u, v) = \frac{1}{B(u, v)} x^{u-1} (1-x)^{v-1} \quad (8)$$

for  $0 < x < 1$ ,  $u > 0$ ,  $v > 0$

with

$$B(u, v) = \int_0^1 x^{u-1} (1-x)^{v-1} dx = \frac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)} \quad (9)$$

For all catchments we assigned the same values of the  $u$ ,  $v$ ,  $p_l$  and  $p_u$  as per Table 1 which has been taken from Merz and Blöschl (2004). If detailed information will be available (for example from catchment attributes or from field studies), the limits and parameters of the Beta distributions for model parameters may be assigned differently from catchment to catchment.

For the evaluation of the calibration and verification efficiencies the entire period of observation (1976–1997) was split into two 11 year periods: the verification period from 1 November 1976 to 31 December 1986 and the calibration period from 1 November 1987 to 31 December 1997. Warm up periods from January to October were used in both cases. Tables 2 to 4 give the model performance of the 320 basins (first line “at-site”) for both the calibration and the verification periods. Figure 2 shows the model performance plotted

**Table 2.** Model efficiency of runoff (*ME*) according to Nash-Sutcliffe for gauged catchments (at-site) and ungauged catchments (various regionalisation procedures) both for the calibration and the verification periods. First value: median *ME* efficiency. Second value: difference of 75% and 25% quantiles of efficiencies, i.e. a measure of scatter. High model performances are associated with large medians and small scatter.

Group	Method	Calibration 1987–1997	Verification 1976–1986
Local	At site	0.72/0.13	0.66/0.20
Mean	Global mean	0.61/0.21	0.56/0.25
Mean	Local mean	0.64/0.18	0.60/0.23
Spatial proximity	Nearest neighbour	0.66/0.18	0.61/0.22
Spatial proximity	Inverse distance weighting	0.66/0.17	0.61/0.21
Spatial proximity	Kriging	0.67/0.16	0.62/0.20
Spatial proximity	Kriging without nested neighbours	0.66/0.16	0.61/0.22
Regression	Global multiple regression	0.60/0.24	0.54/0.28
Regression	Local multiple regression	0.62/0.19	0.58/0.25
Regression	Local georegression	0.65/0.19	0.60/0.22
Similarity	Topography	0.66/0.20	0.61/0.22
Similarity	Geomorphology	0.64/0.19	0.58/0.24
Similarity	Land use	0.65/0.21	0.61/0.25
Similarity	Soils	0.64/0.21	0.59/0.24
Similarity	Geology	0.64/0.20	0.61/0.23
Similarity	Rainfall	0.62/0.21	0.57/0.25
Similarity	Combination	0.67/0.17	0.61/0.21
Similarity	Perfect	0.70/0.14	0.64/0.20

**Table 3.** Volume errors of runoff (*VE*) for gauged catchments (at-site) and ungauged catchments (various regionalisation procedures) both for the calibration and the verification periods. First value: median of *VE* (in %). Second value: difference of 75% and 25% quantiles of *VE* (in %), i.e. a measure of scatter. High model performances are associated with median *VE* close to 0 and small scatter.

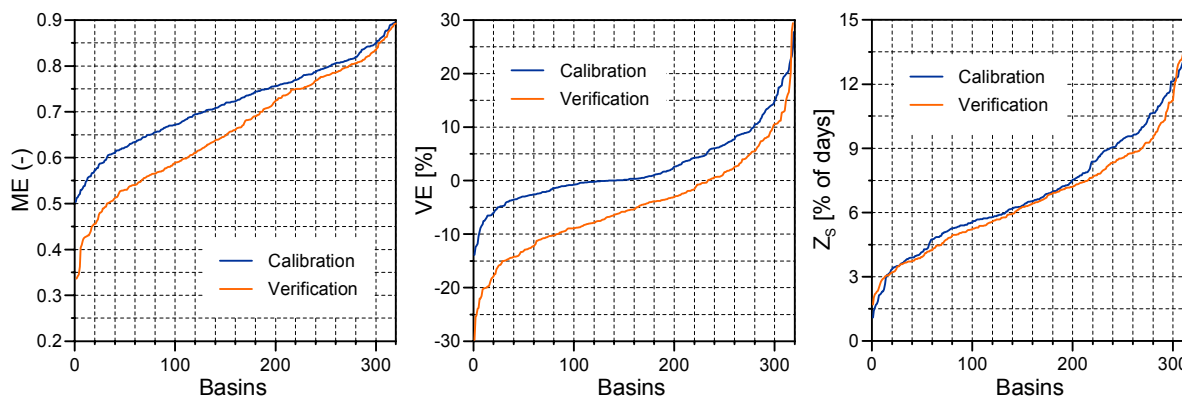
Group	Method	Calibration 1987–1997	Verification 1976–1986
Local	At site	0.3/7.4	−5.3/10.9
Arithmetic Mean	Global mean	−1.3/24.6	−9.2/22.2
Arithmetic Mean	Local mean	−2.2/20.0	−8.6/18.1
Spatial proximity	Nearest neighbour	2.8/18.1	−5.0/17.2
Spatial proximity	Inverse distance weighting	−1.3/18.0	−8.5/16.1
Spatial proximity	Kriging	0.1/16.9	−8.1/16.1
Spatial proximity	Kriging without nested neighbours	0.3/17.6	−8.2/17.1
Regression	Global multiple regression	0.8/27.1	−7.2/25.5
Regression	Local multiple regression	0.8/23.3	−7.3/22.4
Regression	Local georegression	1.0/21.5	−6.9/21.2
Similarity	Topography	1.4/18.5	−6.3/17.1
Similarity	Geomorphology	2.1/20.0	−5.4/20.9
Similarity	Land use	1.5/18.1	−5.5/16.1
Similarity	Soils	2.7/18.0	−4.1/17.8
Similarity	Geology	2.3/18.0	−5.2/17.5
Similarity	Rainfall	2.6/23.7	−5.9/22.1
Similarity	Combination	1.8/20.0	−5.6/17.5
Similarity	Perfect	1.9/9.9	−5.1/13.2

as cumulative distribution functions (*CDFs*). The model efficiency *ME* of runoff is shown on the left, the volume error *VE* of runoff in the centre and the snow cover error on the right. The median of *ME* over the 320 catchments in the calibra-

tion and verification periods is 0.72 and 0.66, respectively. This means that the model performance decreases slightly when moving from calibration to verification. The median of *VE* in the calibration and verification periods is 0.3% and

**Table 4.** Snow cover simulations errors ( $Z_S$ ) for gauged catchments (at-site) and ungauged catchments (various regionalisation procedures) both for the calibration and the verification periods. First value: median of the percentage of days with poor snow cover simulations. Second value: difference of 75% and 25% quantiles, i.e. a measure of scatter. High model performances are associated with small medians and small scatter.

Group	Method	Calibration 1987–1997	Verification 1976–1986
Local	At site	6.55/3.79	6.43/3.39
Mean	Global mean	6.90/4.38	7.10/4.03
Mean	Local mean	6.87/3.96	6.67/3.66
Spatial proximity	Nearest neighbour	6.72/3.74	6.63/3.34
Spatial proximity	Inverse distance weighting	6.77/4.13	6.60/3.39
Spatial proximity	Kriging	6.72/4.23	6.63/3.26
Spatial proximity	Kriging without nested neighbours	6.77/4.18	6.65/3.29
Regression	Global multiple regression	6.97/3.94	6.90/3.86
Regression	Local multiple regression	7.00/3.79	6.55/3.61
Regression	Local georegression	6.97/3.89	6.77/3.29
Similarity	Topography	6.63/3.71	6.40/3.04
Similarity	Geomorphology	6.65/3.89	6.40/3.19
Similarity	Land use	6.60/3.79	6.38/3.36
Similarity	Soils	6.70/3.76	6.67/3.21
Similarity	Geology	6.95/4.11	6.60/3.41
Similarity	Rainfall	6.80/4.18	6.80/3.31
Similarity	Combination	6.65/3.81	6.48/3.11
Similarity	Perfect	6.55/3.99	6.43/3.54

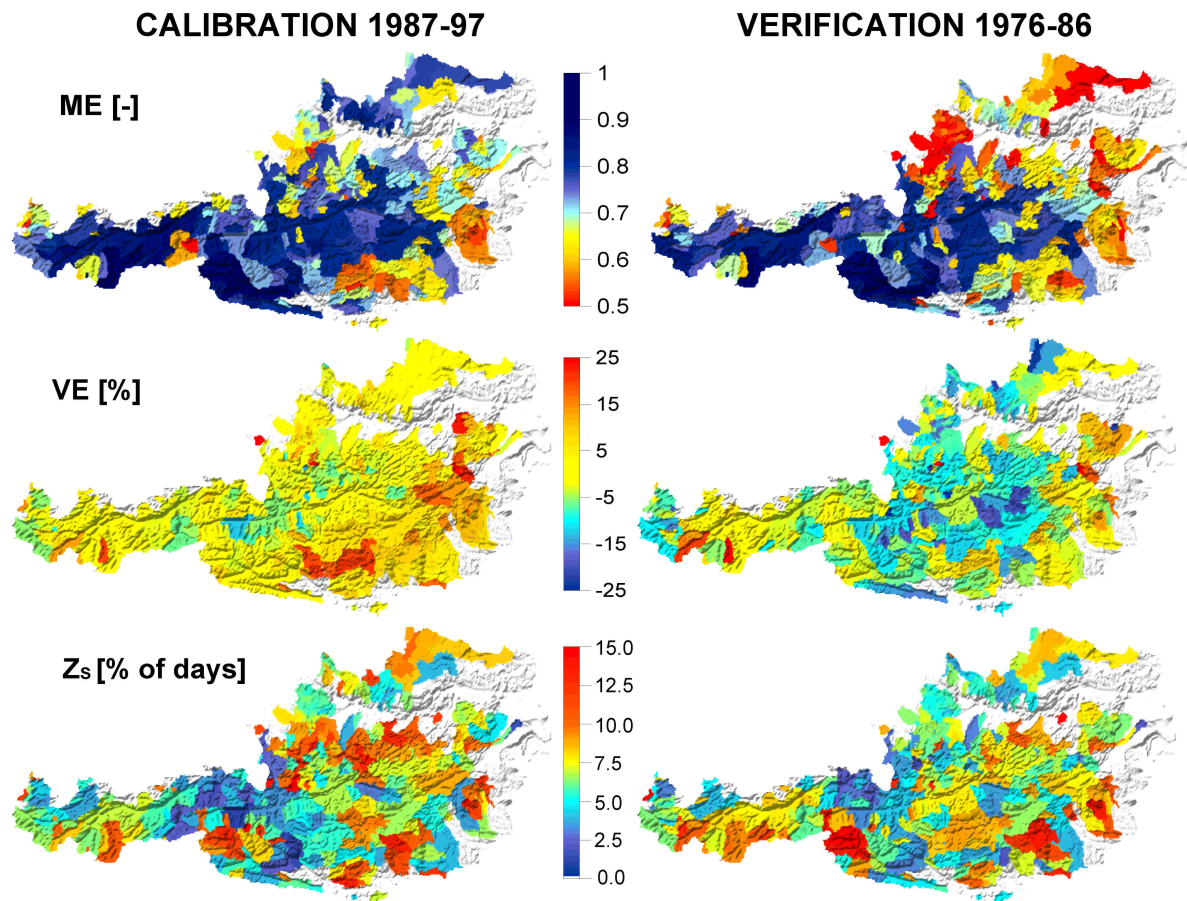


**Fig. 2.** At-site calibration and verification performances: Cumulative distribution functions (CDF) of the model efficiencies of daily runoff ( $ME$ , left), volume errors of runoff ( $VE$ , centre) and percentage of days with poor snow cover simulations ( $Z_S$ , right). 320 basins, calibration (1987–1997) and verification (1976–1986) periods.

–5.3%, indicating that the calibration is essentially unbiased while the verification period does exhibit a small underestimation of runoff. The small underestimation of runoff is related to generally different (drier) climatic conditions observed in verification period 1976–1986. The scatter of the volume error (75%–25% quantile, Table 3) increases somewhat from 7 to 11% which translates into a slightly steeper CDF in Fig. 2 (centre panel) for the case of the verification period. The median of the snow performance measure  $Z_S$  in the calibration and verification periods is 6.5% and 6.4%,

respectively, which indicates that the model performance remains essentially the same.

Overall, the magnitudes of these model efficiencies are similar to results from other regional studies published in the literature (e.g. Seibert, 1999; Perrin et al., 2001). The runoff performances ( $ME$  and  $VE$ ) are somewhat better than those in Merz and Blöschl (2004) even though we used snow data in the objective function, which was not the case in Merz and Blöschl (2004). The differences are 0.03 and 0.05 in terms of  $ME$  for the calibration and verification periods, respectively.



**Fig. 3.** At-site calibration and verification performances: Model efficiencies of daily runoff ( $ME$ , top), volume errors of runoff ( $VE$ , centre) and percentage of days with poor snow cover simulations ( $Z_S$ , bottom). 320 basins, calibration (1987–1997) and verification (1976–1986) periods.

This indicates that the model has been reliably calibrated to the data set.

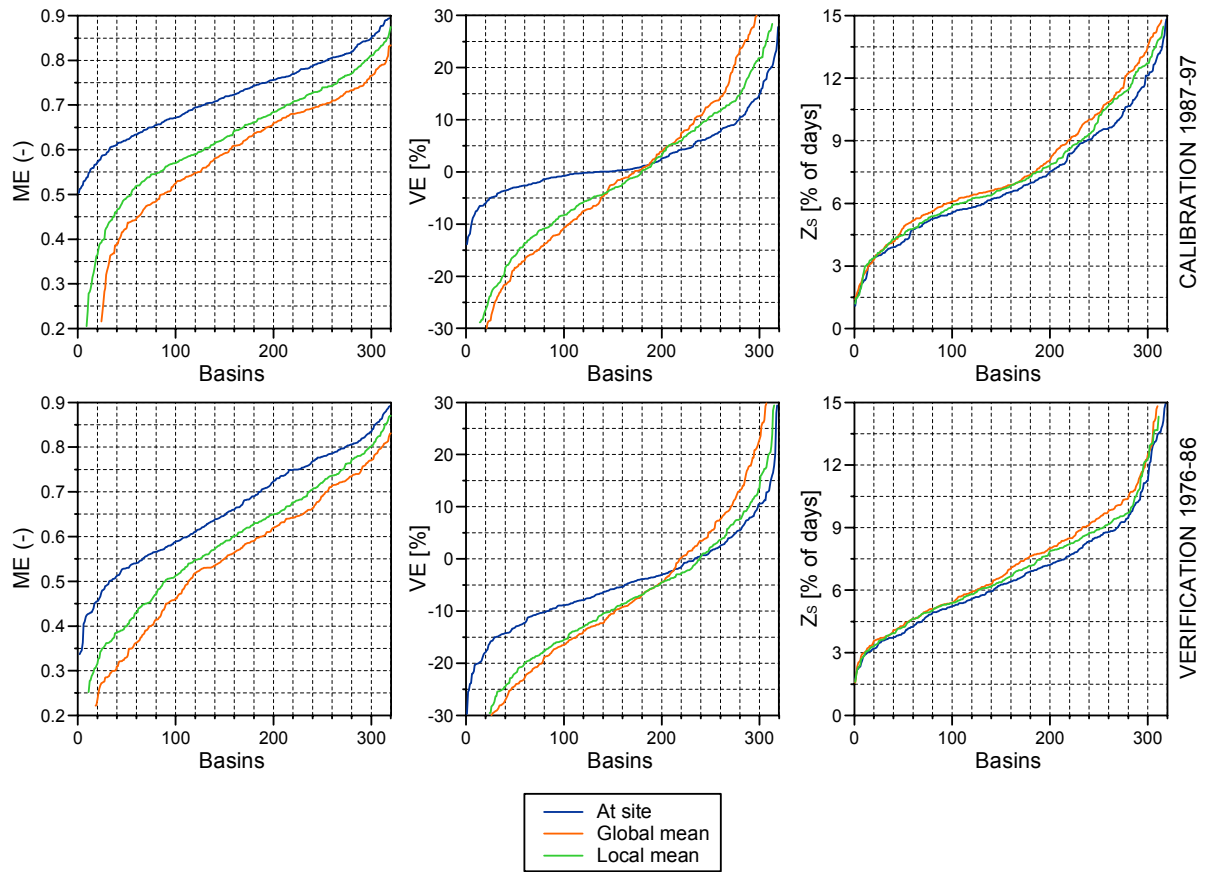
The spatial patterns of the model performances are presented in Fig. 3. The runoff efficiencies  $ME$  and runoff volume errors  $VE$  are shown at the top and the centre, respectively, the snow cover errors  $Z_S$  are shown at the bottom. The left column relates to the calibration period and the right column to the verification period. Figure 3 indicates that there are significant regional differences in the model performance. In the western, alpine parts of Austria the simulation of runoff is significantly better than in the eastern lowlands. The alpine catchments are wetter and snowmelt is more important than in the catchments of the east. It appears that large runoff depths and the presence of snow packs are amenable to accurate runoff simulations.

#### 4 Regionalisation methods

We explored four groups of regionalisation methods. In the first group, we selected each parameter as the arithmetic

mean of all 320 calibrated values in Austria (termed “global mean”) or, alternatively, as the arithmetic mean of a region within a radius of 50 km from the catchment of interest (termed “local mean”). This group of methods assumes that all catchments within the selected radius are similar and differences in the parameter values arise only from random factors.

The second group of regionalisation methods is based on the spatial proximity (or spatial distance) between the catchment of interest and the gauged catchments. The spatial distance between two catchments was measured by the distance of the respective catchment centroids. The methods of this group were the nearest neighbour method where the complete set of model parameters was taken from one donor catchment; the inverse distance weighting where parameters from a number of donor catchments were combined; and the ordinary kriging method. The ordinary kriging method was based on an exponential variogram with a nugget of 10% of the observed variance, a sill equal to the variance, and a range of 60 km. This is consistent with the empirical variograms of



**Fig. 4.** Performance of the spatial averaging group of regionalisation methods: Cumulative distribution functions (CDF) of the model efficiencies of daily runoff (*ME*, left), volume errors of runoff (*VE*, centre) and percentage of days with poor snow cover simulations (*Z<sub>s</sub>*, right). 320 basins, calibration (top) and verification (bottom) periods.

most of the calibrated model parameters. To complement the ordinary kriging method, we also examined ordinary kriging where we left out the immediate upstream and downstream neighbours to assess the effect of nested catchments. We termed this method kriging without nested neighbours.

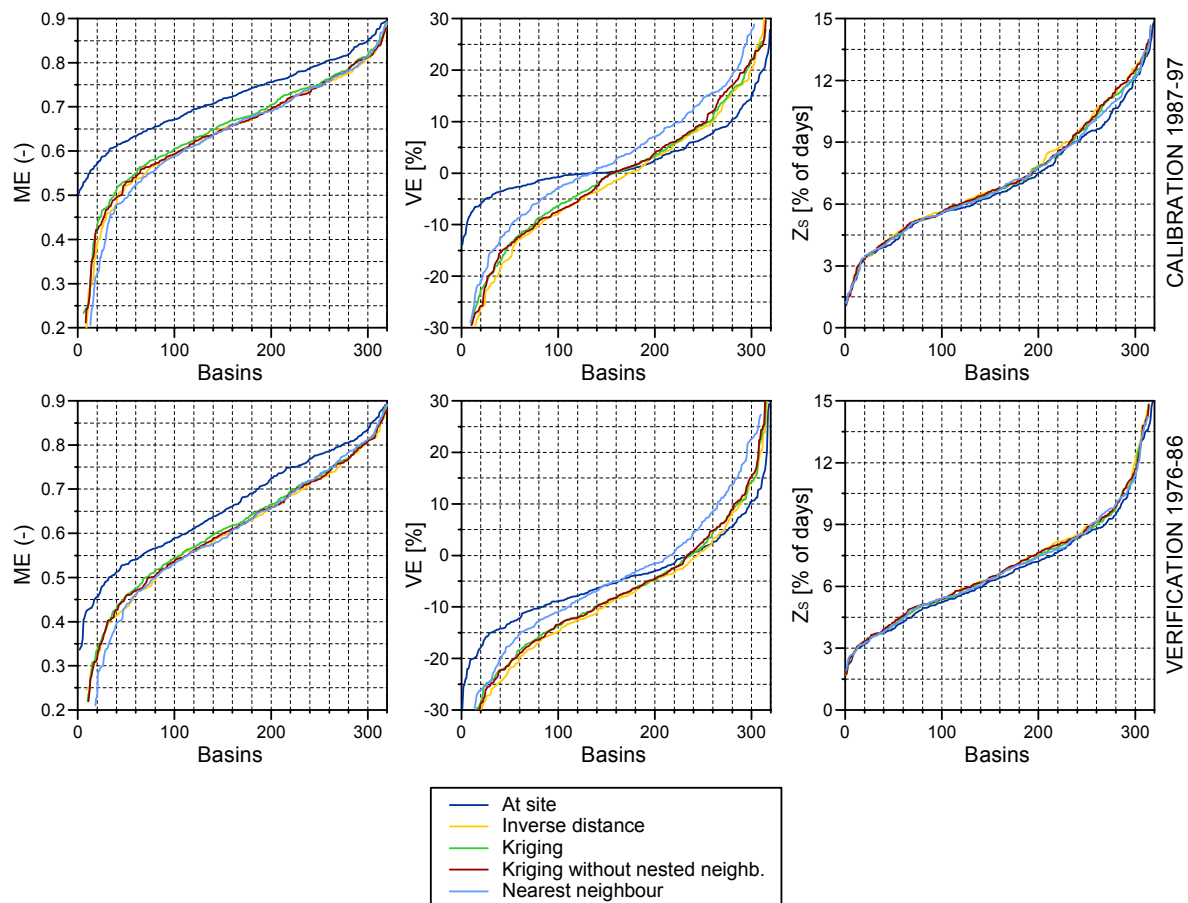
In the third group we estimated each model parameter independently from regressions to catchment attributes. We tested global multiple linear regression, where we included all 320 catchments; local multiple linear regression within a 50 km search radius; and local georegression where we interpolated the residuals of the local multiple regression by ordinary kriging using an exponential semivariogram with 50 km range. In all cases we estimated the regression coefficients by the ordinary least squares method. The number of catchments included in the local multiple regression and the georegression differed regionally. For a 50 km search radius as used here it ranged between 5 and 66 catchments, with an average of 31 catchments. Out of the selected catchment attributes we only used the set of those three attributes that were associated with the largest multiple correlation coefficient for each station and each model parameter. To diagnose

and avoid multicollinearity, we examined the variance inflation factor (Hirsch et al., 1992). If the inflation factor was greater than 10, then this set of three attributes was rejected and the scheme proceeded to the second best correlation. The rationale of this choice is that a large correlation coefficient may be a good indicator of the predictive power of the attributes provided there is no collinearity.

The fourth group of methods is also based on catchment attributes but uses a different regionalisation model structure. The main idea of this group is to find a donor catchment that is most similar in terms of its catchment attributes, and to transpose the complete parameter set to the catchment of interest. Leaving the combination of model parameters unchanged may address some of the problems encountered with the regression approach (Merz and Blöschl, 2004). The donor catchment was selected as the gauged catchment with the smallest similarity index  $\Phi$  (e.g. Burn and Boorman, 1993):

$$\Phi = \sum_{i=1}^k \frac{|X_i^G - X_i^U|}{\Delta X_i}, \quad (10)$$



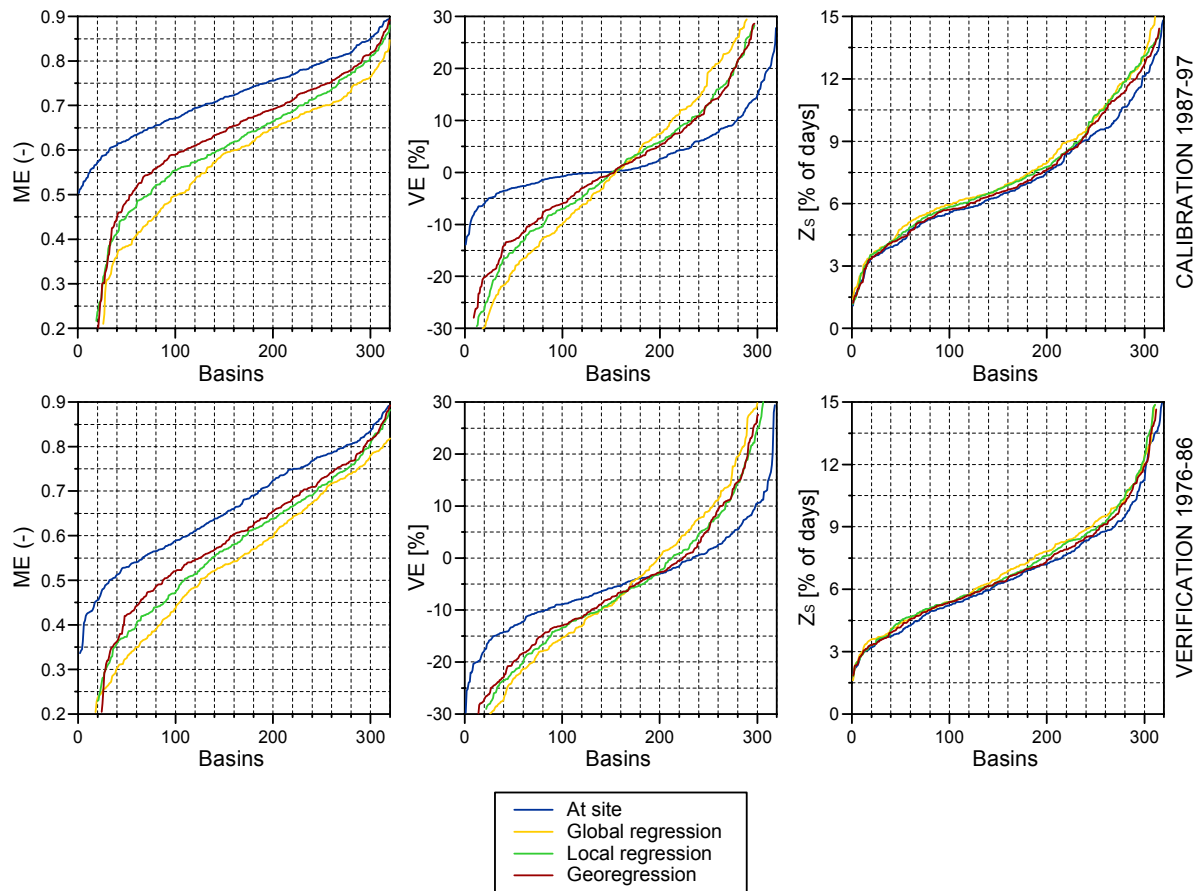


**Fig. 5.** Performance of the spatial proximity group of regionalisation methods: Cumulative distribution functions (CDF) of the model efficiencies of daily runoff (*ME*, left), volume errors of runoff (*VE*, centre) and percentage of days with poor snow cover simulations (*Z<sub>s</sub>*, right). 320 basins, calibration (top) and verification (bottom) periods.

which is defined as the sum of absolute differences of the  $k$  selected physiographic attributes of the gauged ( $X^G$ ) catchment and the (ungauged) catchment of interest ( $X^U$ ), normalized by its range ( $\Delta X$ ). We examined the following combinations of catchment attributes: combinations based on topography (average catchment elevation, slope, topographic index); geomorphology (average stream network density, FEARL index and areal proportion of porous aquifers); land use classes; soils classes; geology classes; rainfall (long-term mean annual precipitation, maximum daily summer and winter precipitation, 1 hourly rainfall intensity); and an a priori defined combination of selected attributes (mean catchment elevation, stream network density, FEARL index and areal proportion of porous aquifers, land use, soils and geologic units). We also tested a diagnostic case termed “perfect”. For the perfect similarity case we transposed the complete parameter set from the donor catchment that was most similar to the catchment of interest in terms of the model parameter values. The similarity was defined by the sum of the absolute differences between the parameter values, normal-

ized by its range similar to Eq. (10). This is a diagnostic case which probes the potential of the catchment model performance that can be achieved with an ideal donor catchment selection. In this study it helps assess the criteria for selecting the catchment attributes used for finding the donor catchment. In a practical application this is not a viable method as the model parameters are of course unknown at the ungauged site of interest. Note that all similarity index based regionalisation methods as well as the geo-regression have not been used in Merz and Blöschl (2004) while the other regionalisation methods have also been examined in Merz and Blöschl (2004).

We examined the predictive accuracy of the regionalisation approaches by jack-knife cross-validation. In this approach, we treated one gauged catchment as ungauged and simulated the water balance dynamics using parameters estimated from regional information only. In a second step, we estimated the model performance by comparing the simulated and observed hydrographs as well as the simulated and observed snow cover. This comparison gave us *ME*, *VE*



**Fig. 6.** Performance of the multiple regression group of regionalisation methods: Cumulative distribution functions (*CDF*) of the model efficiencies of daily runoff (*ME*, left), volume errors of runoff (*VE*, centre) and percentage of days with poor snow cover simulations (*Z<sub>s</sub>*, right). 320 basins, calibration (top) and verification (bottom) periods.

and *Z<sub>s</sub>* model efficiencies. We repeated the analysis for each catchment in turn and calculated the statistics of these error measures for all catchments. The comparison of these error measures with those for the locally calibrated case (here termed “at site”), both for the calibration and verification periods, indicates what decrease of model performance one would have to expect when moving from gauged to ungauged catchments. This decrease we term the spatial loss in model accuracy. The decrease in model performance when moving from the calibration period to the verification period we term the temporal loss in model accuracy.

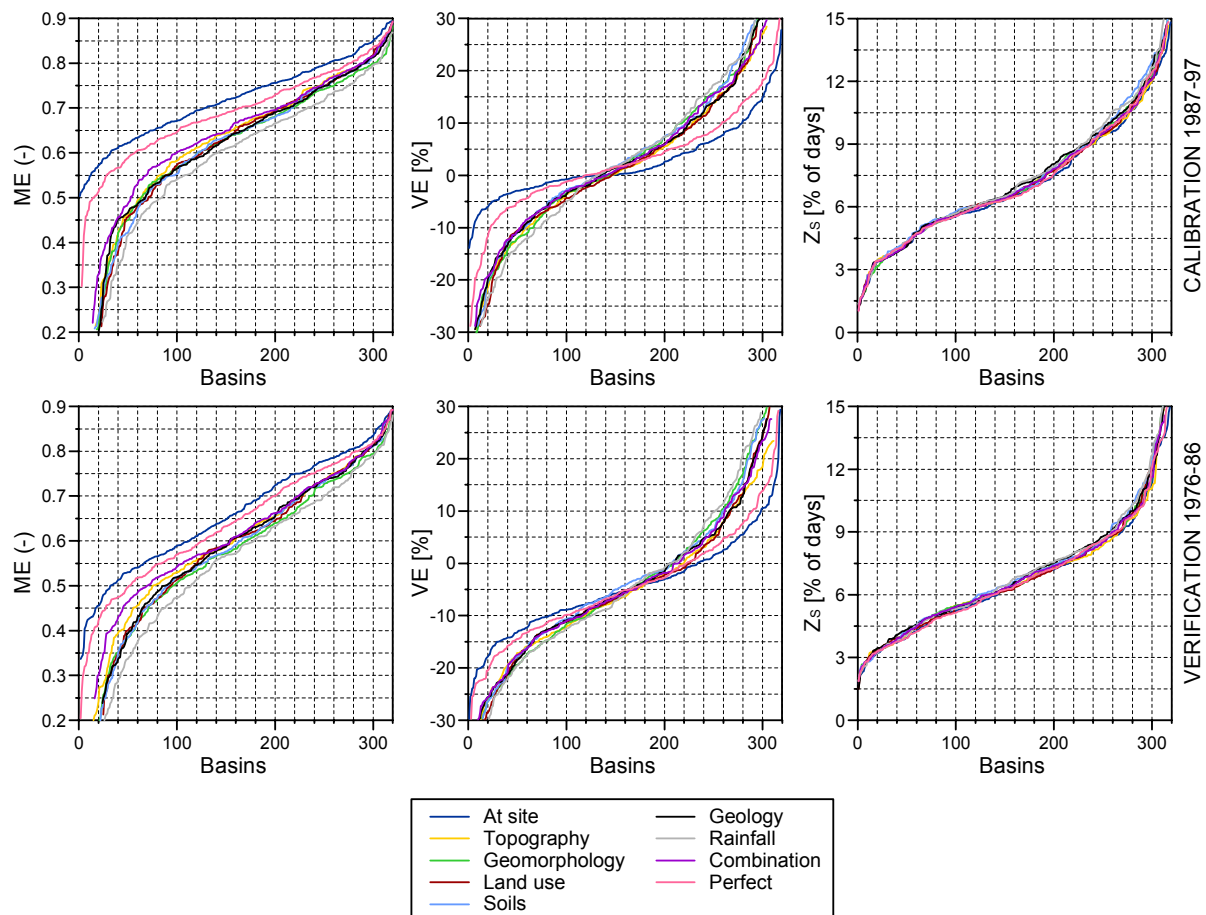
## 5 Performance of regionalisation methods

The performance of the regionalisation methods is presented in terms of their cumulative distribution functions (*CDFs*) in Figs. 4 to 8 and the median and quantile statistics in Tables 2, 3 and 4. For a favourable model performance, the *ME* runoff efficiencies should be large, the *VE* volume errors should be

close to 0 with a small scatter and the *Z<sub>s</sub>* snow cover errors should be small.

Figure 4 shows the performance of the group of methods based on spatial averaging. The left panels show the *ME* runoff efficiencies, the centre panels show the *VE* runoff volume errors and the right panels show the *Z<sub>s</sub>* snow cover errors. The top panels relate to the calibration period and the bottom panels to the verification period. Figure 4 indicates that the global mean method (red line) provides rather poor runoff simulations as compared to the at-site simulations (blue line). The median runoff efficiencies for the global mean method case are *ME*=0.61 (calibration period) as compared to *ME*=0.72 for the at site case (Table 2) and the scatter of the volume errors is much larger (24.6 as compared to 7.4%, Table 3). It is clearly very important to account for differences between catchments, and using the same parameter set for the entire study region is inappropriate for runoff modelling. Using the local mean method slightly improves the efficiencies over the global mean (*ME*=0.64) although the difference is not large. It is interesting that the model





**Fig. 7.** Performance of the similarity index group of regionalisation methods: Cumulative distribution functions (*CDF*) of the model efficiencies of daily runoff (*ME*, left), volume errors of runoff (*VE*, centre) and percentage of days with poor snow cover simulations (*Z<sub>s</sub>*, right). 320 basins, calibration (top) and verification (bottom) periods.

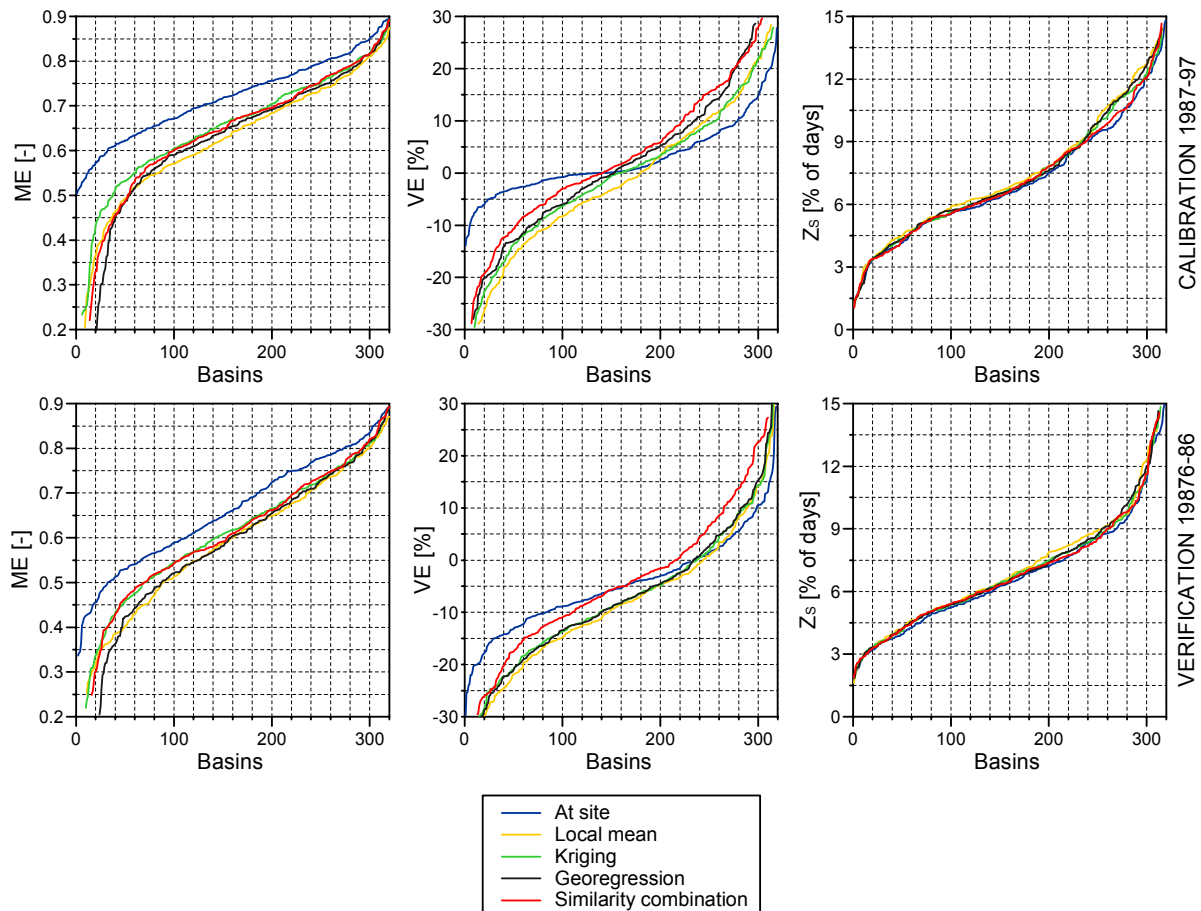
performance in terms of snow cover only decreases slightly when moving from the at-site case to the global mean method case (Table 4).

The *CDFs* of model performances obtained from the regionalisation methods based on spatial proximity are plotted in Fig. 5. The methods only differ very slightly in terms of their runoff performance. The median *ME* runoff efficiency (Table 2) in the case of kriging is  $ME=0.67$  in the calibration period and 0.62 in the verification period. The at-site efficiencies are  $ME=0.72$  and 0.66 which means that the spatial and temporal losses are both about 0.05. The scatter of the *VE* runoff volume error is similar for all spatial proximity methods (about 17% in both periods), which is certainly larger than the scatter of the at-site simulations (7 and 11% in the calibration and verification periods, Table 3). The performance of kriging and kriging without nested neighbours is similar which indicates that the favourable performance of kriging is not only a result of the same portion of the landscape draining into nested catchments. There appear to exist

important similarities of model parameters across catchment boundaries.

Figure 6 shows the results for the multiple regression methods. In this group of methods, the local methods (local regression and georegression) perform better than the global method (global regression). This suggests that it is indeed useful to account for regional differences in the regression equations. The local georegression performs somewhat better than the local regression (e.g.  $ME=0.65$  as compared to 0.62 for the calibration period) which suggests that the spatial correlation of model parameters can enhance the estimates over only using regressions with catchment attributes.

The similarity approach provides an alternative method of using catchment attributes and the results are shown in Fig. 7. The best model performance in terms of *ME* runoff efficiency is provided by the combination similarity measure. The spatial losses are very similar to the kriging approach ( $0.72-0.67=0.05$  in the calibration and  $0.66-0.61=0.05$  in the verification periods). The runoff volume errors, *VE*, are slightly

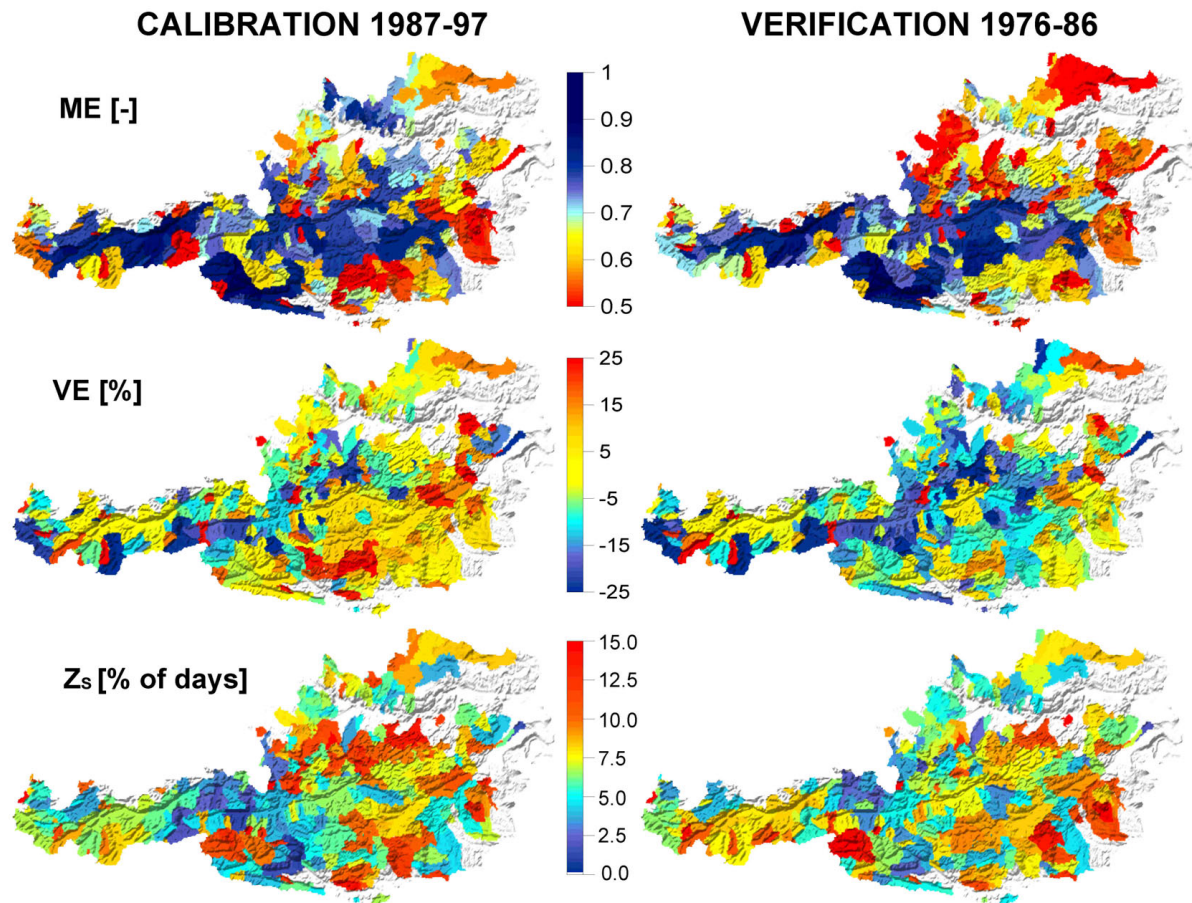


**Fig. 8.** Summary of the performance of the best regionalisation methods of each group: Cumulative distribution functions (CDF) of the model efficiencies of daily runoff ( $ME$ , left), volume errors of runoff ( $VE$ , centre) and percentage of days with poor snow cover simulations ( $Z_s$ , right). 320 basins, calibration (top) and verification (bottom) periods.

larger than for other similarity methods although the difference is not large (e.g. 18% scatter for the land use similarity measure as compared to 20% scatter for the combination similarity measure, both for the calibration period). Again, the snow performances  $Z_s$  are very similar for all the methods. It is interesting to examine the geographical distances between the donor catchments and the catchments of interest. The median distance was similar for all similarity measures and was in the order of 10 km. This suggests that there is significant similarity in the physiographic factors over relatively short distances, which may be one of the reasons for the spatial proximity methods to perform well. The case of the “perfect” similarity index illustrates the model performance when a donor catchment with the most similar model parameters is applied in the water balance simulations. The spatial loss of  $ME$  runoff efficiency is only 0.02 for both the calibration and verification periods which is less than half of the spatial loss of the best regionalisation methods (0.05 in case of both the combination similarity measure and ordinary kriging). The scatter in the volume errors is only 10% as compared to 17%

for the best regionalisation method (Table 3). This indicates that there is indeed potential for improving the criteria for finding donor catchments. For the snow cover errors  $Z_s$  there is, however, very little difference.

The summary of the best regionalisation methods from each group is presented in Fig. 8. The methods are the local mean method, kriging, local georegression, and the combination similarity index approach. The differences between the methods are not large but they do exist. The  $ME$  runoff efficiency shows very little difference for efficiencies of, say,  $ME > 0.7$  but for catchments where the simulated daily runoff fitted poorly to observed values the differences are larger with kriging and the combination similarity index performing best. The  $VE$  runoff volume errors exhibit the most noticeable scatter around the at-site calibration, but it is not possible to ascertain from the CDFs which regionalisation approach produces the smallest scatter. The biases are smallest for the combination similarity method (red line in Fig. 8), at least for the verification period. As indicated in the previous figures, the differences between the regionalisation methods



**Fig. 9.** Performance of the kriging regionalisation method: Model efficiencies of daily runoff (*ME*, top), volume errors of runoff (*VE*, centre) and percentage of days with poor snow cover simulations (*Z<sub>S</sub>*, bottom). 320 basins, calibration (1987–1997) and verification (1976–1986) periods.

are very small in terms of snow performance. For catchments with rather poor snow simulations ( $Z_S > 9\%$ , e.g.) the combination similarity index performs slightly better than other regionalisation approaches.

To examine whether there are spatial patterns in the performance of the regionalisation methods Fig. 9 shows the calibration and verification performances for the case of the kriging regionalisation method. The regional patterns of the *ME* runoff efficiencies (Fig. 9 top) are indeed very similar to those of the at-site calibration and verification (Fig. 3 top). There appears to exist, however, more spatial scatter, which is mainly due to a number of small catchments in the central alpine parts of Austria, where the regionalisation performed poorer. While the median spatial loss in model performance over all catchments was 0.05 in the calibration period it is larger for catchments with areas of less than  $100 \text{ km}^2$  (about 0.09). This indicates that in small catchments the peculiarities in runoff forming conditions are more difficult to capture than in larger catchments where always some sort of averaging takes place.

The spatial patterns of the *VE* runoff volume errors (Fig. 9), again, exhibit larger scatter than those of the at-site case (Fig. 3). This is consistent with the larger scatter (75%–25% quantiles) indicated in Table 3 and the steeper *CDF* shown in Fig. 5. The difference is particularly large in the high alpine parts of Austria, which is consistent with the large scatter in *ME* as shown in the top panels of Fig. 9. The regional snow cover errors also show somewhat larger scatter in the regionalised case as compared to the at-site simulations.

## 6 Discussion and conclusions

The results indicate that two regionalisation methods perform best. The first is a kriging approach where the model parameters are regionalised independently from each other based on their spatial correlation. The second is a similarity approach where the complete set of model parameters is transposed from a donor catchment that is most similar in

terms of its physiographic attributes (mean catchment elevation, stream network density, lake index, areal proportion of porous aquifers, land use, soils and geology). The first result is consistent with Merz and Blöschl (2004) who indicated that spatial proximity may be a better similarity measure for transposing catchment model parameters in space than physiographic catchment attributes. We improved the model structure over that used in Merz and Blöschl (2004) by allowing for elevation zones and we enhanced the parameter estimation by using snow data in addition to runoff but the finding of the favourable performance of kriging remains the same. Similar to Merz and Blöschl (2004), there is only a slight decrease in model performance when leaving out the immediate (nested) neighbours in the regionalisations. This suggests that the favourable performance is not only a result of the same portion of the landscape draining into nested catchments. There appear to exist important similarities of model parameters across catchment boundaries. It is likely that these similarities are related to real hydrological controls that vary smoothly in space. For a number of catchments the regionalisation does perform poorly with efficiencies one would not use in practical applications. This is particularly the case in the high alpine areas where the spatial hydrologic variability is particularly large. Also, in some low land catchments the runoff model does not seem to represent the runoff dynamics very well. Vandewiele and Elias (1995) have pointed to similar issues, which they traced back to both spatial hydrologic variability and poor data quality.

The second result of the favourable performance of the similarity approach using physiographic catchment attributes is interesting in the light of the relatively poor performance of the regression approach found both in Merz and Blöschl (2004) and in this paper. One of the advantages of the similarity approach may be that the complete set of model parameters is transposed from a donor catchment. This is consistent with the findings of Kokkonen et al. (2003, p. 2219), who concluded that “when there is a reason to believe that, in the sense of hydrological behaviour, a gauged catchment resembles the ungauged catchment, then it may be worthwhile to adopt the entire set of calibrated parameters from the gauged catchment instead of deriving quantitative relationships between catchment descriptors and model parameters”. The other advantage of the similarity approach over the regression method as used in this paper is that it does not make the assumption of linearity. The main reason of using linear regression models is that of convenience although the underlying hydrological relationships are unlikely linear in nature. In our study, we tested various combinations of similarity indices. The favourable performance of the diagnostic index termed “perfect” suggests that there still exists potential for improving the regionalisation methods by identifying more relevant physiographic controls.

Overall, the model performance is similar to that of other regionalisation studies in the literature. Seibert (1999) reported of a median loss in *ME* runoff efficiency from 0.81

to 0.79 when moving from calibrated parameters to regionalised parameters for the same set of 11 catchments, but a decrease to 0.67 for a separate set of 7 catchments. Beldring et al. (2002) found median *ME* of 0.68 for both a set of 141 gauged catchments and a set of 43 catchments treated as ungauged although for 20% of the catchments belonging to the second set the efficiencies were less than 0.3. As compared to Merz and Blöschl (2004), the *ME* model performances increased by between 0.07 and 0.10 depending on the regionalisation method. This is mainly due to the improved model structure of allowing for elevation zones. However, the uncertainty is still large. As pointed out by Blöschl (2005), site visits involving a field assessment of catchment behaviour may assist in improving the model performance beyond what can be realistically achieved from catchment attributes that are available at the regional scale.

The comparisons of the regionalisation methods indicated that the overall snow performance is almost insensitive to the choice of method. Detailed analysis of snow model efficiency in particular catchments revealed three aspects that affect the similar snow model performance. First finding is that in many catchments there are differences in snow simulation performed by different regionalisation approaches, but overall snow performance measure (median over 320 catchment) does not make a distinction between them. Next obvious issue is that out of the 11 calibrated model parameters it is only the degree day factor and the snow correction factor that will affect snow simulations. The three other snow model parameters were preset. However, the lack of sensitivity may also be related to the snow data and the spatial snow interpolation. In this study, point snow depth measurements have been spatially interpolated and the point data may not be very representative of the catchment snow cover. One possibility of improving the spatial representativeness is the use of satellite snow cover data (e.g. Grayson et al., 2002) which will be pursued in further research.

Another interesting issue for further research is to test the model performance if more data sets are left out in order to verify the predictive accuracy of different regionalisation approaches. This would provide some indication of how the regionalisation approach would work for regions with smaller datasets. For the greater discrimination between different regionalisation methods we plan to apply additional measures of model performance such as the statistics of annual and seasonal daily peak or low flows, snow similarity measures based on patterns comparison etc.

**Acknowledgement.** The authors would like to thank the Marie Curie Fellowship of the European Community programme HUMAN POTENTIAL under contract number HPMF-CT-2002-01872, and the Austrian Academy of Sciences, project HÖ 18, for financial support. We would also like to thank the Austrian Hydrographic Service (HZB) for providing the hydrographic data. We are grateful also to the two reviewers, M. Pfaundler and J. Szolgay, and to the editor, P. Molnar, for their thoughtful comments and suggestions.



Edited by: P. Molnar

## References

- Beldring, S., Roald, L. A., and Voksø, A.: Avrenningskart for Norge (Runoff map for Norway, in Norwegian), Norwegian Water and Energy Directorate, Report No. 2, Oslo, Norway, 2002.
- Bergström, S.: Development and application of a conceptual runoff model for Scandinavian catchments, Dept. of Water Resour. Engineering, Lund Inst of Technol./Univ. of Lund, Bull. Ser. A, No. 52, 1976.
- Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, *Hydrol. Sci. Bull.*, 24, 43–69, 1979.
- Blöschl, G.: Rainfall-runoff modelling of ungauged catchments, *Encyclopedia of Hydrological Sciences*, John Wiley & Sons, in press, 2005.
- Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling – a review, *Hydrol. Proc.*, 9, 251–290, 1995.
- Burn, D. H. and Boorman, D. B.: Estimation of hydrological parameters at ungauged catchments, *J. Hydrol.*, 143, 429–454, 1993.
- Campbell, E. P. and Bates, B. C.: Regionalisation of rainfall-runoff model parameters using Markov Chain Monte Carlo samples, *Water Res. R.*, 37, 3, 731–739, 2001.
- Duan, Q., Sorooshian, S., and Gupta, V. K.: Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Res. R.*, 28, 1015–1031, doi:10.1029/91WR02985, 1992.
- Ecker, R., Kallian, R., and Steinnocher, K.: Fernerkundungsdaten für die Planung eines Mobilfunknetzes, *Österr. Zeitschr. f. Vermessung und Geoinformation*, 83, 14–25, 1995.
- Fernandez, W., Vogel, R. M., and Sankarasubramanian, A.: Regional calibration of a watershed model, *Hydro. Sci. J.*, 45, 5, 689–707, 2000.
- Flood Estimation Handbook: Institute of Hydrology, Wallingford, 1999.
- Geologische Bundesanstalt: Metallogenetische Karte 1:500.000 auf CD-ROM, GBA, FA-ADV, Wien, 1998.
- Grayson, R., Blöschl, G., Western, A., and McMahon, T.: Advances in the use of observed spatial patterns of catchment hydrological response, *Adv. Wat. Res.*, 25, 1313–1334, 2002.
- Hlavčová, K., Szolgay, J., Čistý, M., Kohnová, S., and Kalaš, M.: Estimation of mean monthly flows in small ungauged catchments, *Slovak Journal of Civil Engineering*, Vol. VIII, 21–29, 2000.
- Hirsch, R. M., Helsel, D. R., Cohn, T. A., and Gilroy, E. J.: Statistical analysis of hydrological data, in: *Handbook of Hydrology*, edited by: Maidment, R., McGraw-Hill, New York, 17.1–17.55, 1992.
- HZB: Hydrologisches Jahrbuch von Österreich 1997, Hydrographisches Zentralbüro im BMLF, Band 105, Wien, 2000.
- Kokkonen, T. S., Jakeman, A. J., Young, P. C., and Koivusalo, H. J.: Predicting daily flows in ungauged catchments: model regionalisation from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina, *Hydrol. Proc.*, 17, 2219–2238, doi:10.1002/hyp.1329, 2003.
- Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, *J. Hydrol.*, 201, 272–288, doi:10.1016/S0022-1694(97)00041-3, 1997.
- Merz, R. and Blöschl, G.: Regionalisation of catchment model parameters, *J. Hydrol.*, 287, 95–123, doi:10.1016/j.jhydrol.2003.09.028, 2004.
- Mészáros, I., Miklánec, P., and Parajka, J.: Solar energy income modelling in mountainous areas, in: *ERB and NEFRIEND Proj.5 Conf. Interdisciplinary Approaches in Small Catchment Hydrology: Monitoring and Research*, edited by: Holko, L., Miklánec, P., Parajka, J., and Kostka, Z., Slovak NC IHP UNESCO/UH SAV, Bratislava, Slovakia, 127–135, 2002.
- ÖBG: Bodenaufnahmesysteme in Österreich, *Mitteilungen der Österreichischen Bodenkundlichen Gesellschaft*, Heft 62, 2001.
- Parajka, J., Merz, R., and Blöschl, G.: Estimation of daily potential evapotranspiration for regional water balance modeling in Austria, in: *11th International Poster Day and Institute of Hydrology Open Day “Transport of Water, Chemicals and Energy in the Soil – Crop Canopy – Atmosphere System”*, Slovak Academy of Sciences, Bratislava, 299–306, 2003.
- Pebesma, E. J.: Gstat user’s manual, Dept. of Physical Geography, Utrecht University, Utrecht, The Netherlands, 2001.
- Peel, M. C., Chiew, F. H. S., Western, A. W., and McMahon, T. A.: Extension of unimpaired monthly streamflow data and regionalisation of parameter values to estimate streamflow in ungauged catchments. Report prepared for the National Land and Water Resources Audit, in: *Australian Natural Resources Atlas website*: <http://audit.ea.gov.au/anra/water/docs/national/Streamflow/Streamflow.pdf>, 2000.
- Perrin, C., Michel, C., and Andréassian, V.: Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, *J. Hydrol.*, 242, 275–301, 2001.
- Post, D. A. and Jakeman, A. J.: Relationships between catchment attributes and hydrological response characteristics in small Australian mountain ash catchments, *Hydrol. Proc.*, 10, 877–892, 1996.
- Post, D. A. and Jakeman, A. J.: Predicting the daily streamflow of ungauged catchments in SE Australia by regionalising the parameters of a lumped conceptual rainfall-runoff model, *Ecol. Model.*, 123, 91–104, 1999.
- Sefton, C. E. M. and Howart, S. M.: Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales, *J. Hydrol.*, 211, 1–16, 1998.
- Seibert, J.: Regionalisation of parameters for a conceptual rainfall-runoff model, *Agr. For. Met.*, 98–99, 279–293, 1999.
- Shu, Ch. and Burn, D. H.: Spatial patterns of homogeneous pooling groups for flood frequency analysis, *Hydro. Sci. J.*, 48, 4, 601–618, 2003.
- Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., McDonnell, J. J., Mendiola, E. M., O’Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E.: IAHS decade on predictions in ungauged basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, *Hydro. Sci. J.*, 48, 6, 867–880, 2003.
- Szolgay, J., Hlavčová, K., Kohnová, S., and Daníhlík, R.: Estimating regional parameters of a monthly balance model, *J. Hydrol. Hydromech.*, 51, 4, 256–273, 2003.
- Vandewiele, G. L. and Elias, A.: Monthly water balance of ungauged catchments obtained by geographical regionalisation, *J. Hydrol.*, 170, 277–291, 1995.